DOCUMENT RESUME

ED 178 416 SO 012 076

AUTHOR	Cort , H. R., Jr.; And Others
TITLE	Approaches to Further Evaluation of Man: A Course of Study. A Final Report.
INSTITUTION	Washington School of Psychiatry, Washington, D.C.
SPONS AGENCY	National Science Foundation, Washington, D.C.
PUB DATE	19 Feb 71
NOTE	78p.; For related documents, see SO 012 048, ED 045 461 and ED 065 458; Best copy available. The summary and concluding comments are cut short due to missing pages
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Course Evaluation; *Curriculum Development; *Curriculum Evaluation; Educational Objectives; Elementary Education; Evaluation Criteria; Evaluation Methods; Experimental Curriculum; *Social Studies; Social Studies Units
IDENTIFIERS	Man A Course of Study

ABSTRACT

The paper reports on an investigation of evaluation alternatives of the social studies curriculum project "Man: A Course of Study" (MACOS). It is presented in eight sections. Section I describes the task of the study and offers a view of the stages of curriculum development projects. Section II describes methods employed in the study. Section III summarizes findings from the field visits relating to the environment in which the project was used, dissemination, functions of the course, the value of the course for different ethnic groups and grade levels, teacher preparation, impact on the institution, and previous evaluation. Section IV presents further examples and trends, including teachers' perceptions of MACOS goals, how teachers evaluate student progress, and context considerations. Section V outlines 15 characteristics of the MACOS curriculum, while Section VI cites six issues or categories to be considered in further evaluation. Section VII presents three specific approaches: absolute evaluation, comparative evaluation, and systemic evaluation. It analyzes the advantages and limitations of each in relation to general design requirements, specific design and methodology, implementation mechanisms, and objectives and payoffs. Section VIII summarizes and offers conclusions, including that it is possible to evaluate MACOS using available methods and instruments, and that there is no one measure of the effectiveness of the program. Further conclusions are cut short due to missing pages. (CK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED178416

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Mary Louise Charles, NSF

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

APPROACHES TO FURTHER EVALUATION OF MAN: A COURSE OF STUDY

BEST COPY AVAILABLE

A Final Report
by

H. R. Cort, Jr.
Naomi H. Henderson
Cheryl Jones

to

THE NATIONAL SCIENCE FOUNDATION

February 19, 1971

SP 012 076

FOR INTERNAL USE ONLY

Educational Studies Department
Washington School of Psychiatry
1610 New Hampshire Avenue, N.W.
Washington, D.C., 20009

TABLE OF CONTENTS

		Page
Introduction		i
Section I	A Suggested Conceptual Framework	1
Section II	Description of ESD's Procedures	6
Section III	Brief Summary of Information Gathered	8
Section IV	Further Examples and Trends From Interviews	13
Section V	Some Critical Features of the MACOS Curriculum	26
Section VI	Some Considerations Concerning Further Evaluation of MACOS	27
Section VII	Alternative Strategies or Approaches	43
Section VIII	Summary and Concluding Comments	72
Appendix		76

INTRODUCTION

This is the final report of an investigation of further evaluation alternatives for the social studies curriculum Man: A Course of Study.* The purpose of the study was not to specify an approach, but to clarify alternatives and bases for choosing among them. As a preliminary step in developing suggestions, field visits were made to seven MACOS Regional Centers,** and to several school districts and schools in each region, to obtain information about conditions of use of the curriculum. The observations made and information and impressions obtained entered strongly into our ideas about further evaluation needs and strategies.

In the course of preparing this report we found ourselves rethinking most of our preconceptions and assumptions about further evaluation of the MACOS package, especially summative evaluation. The question of feasibility of alternative evaluation approaches became integrally enmeshed with the prior question: evaluation to what end, for what purpose? The reader will find that we tried consistently to address this question on a gross level, but that we have not been able to provide a clear analytic framework for answering it.

The structure of this report is as follows. In Section I the

---

* This report was prepared under NSF Grant No. W005707.

** One center had recently been terminated.

task of the study and a general view of the processes of curriculum development projects are described. Section II summarizes the general methods employed in this study. Section III gives a brief summary outline of "findings" from field visits. In Section IV we present further examples and discussion of findings, and Section V provides a summary list of characteristics of the curriculum that appeared important or significant to various respondents in the field. In Section VI we then discuss six categories of evaluation questions or concerns related to MACOS. In most categories we note characteristics of the curriculum that seem especially germane to the design of further evaluation studies. We then synthesize in the next section the analyses made in Section VI into three major kinds of further evaluation that might be undertaken. In Section VII we discuss design requirements for, methods and potential values of absolute, comparative, and systemic evaluations that could be undertaken. We finally provide a summary statement, in Section VIII, of issues, alternatives and recommendations.

It will be apparent to the reader that we found our task complex. In an effort to avoid vigorous simplistic suggestions we constantly risked paralysis from attention to a multitude of issues, uncertainties and possibilities. The net result of efforts to achieve a balance will undoubtedly strike many as erring in one direction or the other. For ourselves, we believe that erring in the direction of complexity is the better, if aesthetically less satisfying, error.

I

A SUGGESTED CONCEPTUAL FRAMEWORK

The National Science Foundation expressed its interest in funding a feasibility study to determine what some possible next steps, in terms of continuing evaluation of MACOS, might be. A thorough and excellent formative evaluation had been completed.[1] Discussions with NSF did not provide us with any clear answers to the question, "What are the NSF's further interests or goals with MACOS?" Therefore, we suggested that we try to depict some alternative strategies for continued evaluation based on what we could determine to be the needs, present conditions, and concerns of various people in the field.

Several questions came to mind immediately. What has happened with the development and dissemination of the curriculum to date? What is the present state of its utilization? How effective is MACOS compared with other social studies curricula? And so on. We felt that by talking with classroom teachers, social studies curriculum specialists within school systems, administrators, regional directors and disseminators, and members of the Education Development Center, we could help focus on these and other questions. We hoped such an approach might help outline some priorities for further study. In an effort to do this we proposed a series of interviews in each of the regions established for dissemination of MACOS. Using the data gathered we shall, in this report, suggest some evaluation issues and strategies for NSF's consideration.

---

[1] Hanley, Janet P., et al. Curiosity/Competence/Community: An Evaluation of Man: A Course of Study. Cambridge, Mass.: Social Studies Curriculum Program, Education Development Center, Inc., 1970. Vols. i & ii.(Preliminary copy.)

A review of MACOS materials, especially the most recent Seminars for Teachers,[2] helped to clarify the potential multiple functions the new curriculum might have. It seemed that a consideration of evaluation issues with respect to the various possible functions of the curriculum might provide a way of organizing and integrating the concerns of emerging educational evaluation theory with the realities of MACOS research, development, and utilization. Our hope then, was that we could clarify, at least presumptively, the manifold number of issues involved in MACOS as well as suggesting some priorities for further study.

There are four broad functions a new curriculum may perform:

1. Education of children
2. Development and enhancement of pedagogical skills and orientation of teachers
3. Change in institutions
4. Training and development of prospective teachers

There is nothing new about these areas of effect, impact, or concern. Most of them are discussed in one form or another, for example, by Grobman.[3]

Patently, the purpose of a new curriculum such as MACOS is to improve the educational growth and development of students. Equally obviously the educational enterprise is a manifold affair, fraught with uncertainties about what causes what, or what maximizes what. Under such circumstances, there are legitimately different strategies and options for improvement. Furthermore, the same device may in fact bear on different dimensions of the educational enterprise, depending on how and where it is used. A

---

[2] Man: A Course of Study. Seminars for Teachers. Education Development Center, Inc., 1970.

[3] Grobman, Hulda. Evaluation Activities of Curriculum Projects. AERA Monograph Series of Curriculum Evaluation, No.2. Chicago: Rand McNally, 1968.

curriculum package, then, may perform some or all of the functions listed above. It is a matter of considerable interest to try to explicate rules or propositions for which ones, when, where, how, etc.

If one is going to design a new curriculum, he obviously has the intention of producing something that will have one or more of the functions suggested above. But between the design intention and its accomplishment of one or more of the curriculum functions are a number of linking or enabling functions. E.g.,

1. Development and testing
2. Dissemination
3. Preparation and installation[4]
4. Implementation
5. Maintenance and support
6. Expansion, modification or re-direction
7. Monitoring and evaluation[5]

Suppose it is the designer's intention to produce a curriculum package that will perform particular educational functions for or with elementary school children (in the case of MACOS, for example, to develop hypothesis forming skills, to speak glibly for the moment.) The design and development of certain materials and procedures constitutes a transformation of that intention into a mechanism or apparatus for performing that specific educational function. Tests of the curriculum are used to determine whether, and presumably under what conditions, the package does perform it.

---

[4] Preparation here includes orientation and training of teachers and others prior to or at the start of the curriculum; installation includes purchasing, scheduling, etc. Jointly they designate the activities necessary or undertaken within a school or school system to put the new curriculum into effect.

[5] This should be a parallel function, not a serial or partially serial one like the others.

One evaluation issue at this point concerns the identification of intended functions, their performance characteristics, and the limiting assumptions concerning the performance characteristics (e.g., what skills of teachers are assumed? what kinds of students? what organizational arrangements in schools? etc.). A second evaluation issue at this point concerns the design and development method or procedure. What did it take to accomplish the transformations? What are the implications for further curriculum project funding decisions? Can design and development components be related to costs in a generalizable sense?[6]

The curriculum or mechanism intended to generate the designated function is then disseminated, that is, brought to the attention of prospective users. In this transformation the intended function may be retained, modified or lost, and/or other functions and potential functions added. Some number of prospective users will adopt the new curriculum over a given period of time. Processes of preparation, installation, and eventually implementation thus commence. Evaluation issues to that point include:

1. what are the operating characteristics of different dissemination mechanisms? are some more efficient than others? how? with what kinds of users? with respect to what kinds of functions?

2. is there a relationship between different kinds of dissemination

---

[6] We are trying here to get at the problem of evaluating what kinds of approaches, with what kinds of variations, yield what kinds of products (curriculum packages, for example) under what conditions. We think this sort of analysis is related to further decision-making about the need for public support of curriculum design and development.

mechanisms and the further transformation of intended functions?[7]

The new curriculum is adopted and implemented, tentatively or permanently. At this stage there is the final set of transformations leading to resultant functions. With respect to the function dealing with education of children, the evaluation issues center around the traditional problems of measuring effectiveness, or more specifically, differential effectiveness. That is, who learns what, under what conditions? Are there unexpected results? Are the designer's intended functions realized? What support conditions turn out to be needed? What further modifications are appropriate? And so on.

This highly simplistic discussion is intended only to suggest that:

1. there are different evaluation issues at different stages of a curriculum development project;

2. the issues are separable with respect to certain decision problems for certain agencies; but data bearing on issues at one point may need to be carried forward to apply at another point;

3. the transformations that may occur are only partly predictable, though eventually they are completely (potentially) knowable.

We shall return to these themes. They are much broader than the concept of summative evaluation. This brief discussion of an overall process indicates that in approaching our task we have tried to consider a broad spectrum of congruences, contingencies, and decisions.

---

[7] The decision problems here center around the development and support of dissemination mechanisms. Is there a rational basis for the allocation of funds to different dissemination designs? What are design parameters?

II

DESCRIPTION OF ESD's PROCEDURES

Our first steps were to make ourselves more familiar with the MACOS package and its history. We did this through consultation with members of the Innovation Team in the District of Columbia who had been teachers, disseminators, and trainers for MACOS; through the review of the curriculum materials and documents about the program; and through discussion with the MACOS evaluation staff at Education Development Center. We then prepared questionnaires to be answered in interviews with regional center directors and people associated with regional centers. The questionnaires were also to be used with school administrators, curriculum supervisors, and teachers. The basic intentions of the questions were to provide information about what observers saw as the purposes, both actual and potential, of MACOS; what methods were being used to assess or evaluate the program; what was being done to disseminate the program, install it, and support it; and what seemed to be some of the crucial factors involved in each of these activities.

The Educational Studies Department Team visited seven regional centers, twenty-four school districts, and thirty-six schools within districts. School districts visited were selected in discussion with center directors, and final choice was based on interest and practicality. The schools in no sense represent a sample, but it is perfectly clear to the Team that they covered a wide variety of environments and educational conditions. We were able to compose a detailed compendium of responses to various questions, and thus to identify potentially important variables, based either on direct statements gleaned from the talks we had with many people, or on impressions derived from the overall discussions. In addition, we have done an extensive

perusal of the evaluation literature in general, and of literature concerned with social studies and educational change in particular. In Appendix A there is a table summarizing the number and characteristics of teachers interviewed, and numbers of locations visited.

III

BRIEF SUMMARY OF INFORMATION GATHERED

The following is a condensed, selective summary of information and impressions gleaned from our field visits. It is not intended to be a definitive or conclusive set of statements.

I. Environment in which curriculum is used or may be used

A. Target groups

Small cities--largest group studied--offer least amount of red tape and hierarchy to deal with; key decision makers stand out.

Suburban areas--usually can afford program, often their schools already exposed to progressive or innovative materials and pedagogies.

Rural areas--often can't afford materials; often conservative; hard to follow up.

Large metropolitan areas--most difficult to work with because of bureaucracy, difficulty in pinpointing key decision makers. Politics often cause decision makers to be reluctant to make a commitment to the program.

B. Political climate runs the gamut from hard core conservative to liberal-progressive seeking all kinds of innovative curricula.

C. Financial resources

Sometimes materials provided through special innovative funding programs, or on an experimental or pilot basis.

Cost a problem to many small systems which simply find it prohibitive.

Same cost problem when an individual school wants the materials but is not supported by the system.

D. Educational climate (see remarks under Target groups)

Some systems more interested in innovative curricula than others; interest cuts across urban/rural classifications.

Often reflects political climate.

Progressiveness of key decision makers influences educational climate.

II. Dissemination

A. No one process across regions or within regions.

B. Differs from system to system, but the more bureaucratic and larger the system, the harder the dissemination process, except where key decision makers seek out MACOS.

C. Financial resources of system or school again play a large role.

D. Political climate an influence; the more conservative the community the less receptive it may be to MACOS.

III. Support

A. No pattern across regions. Ranges from built-in support system through summer institute training of social studies curriculum consultant for each system, to no support system at all.

B. In most cases support is given on an "as can" basis.

C. Problem for regional directors is the lack of funds and staff to adequately follow up teachers. Regions quite large.

D. Many teachers feel more follow-up would be helpful, even those who have taught the program for as many as four years. New teachers to the program need beginning skills. Older teachers to the program need revitalizing.

IV. Functions

A. Target groups

Children

Most people, from teachers to regional directors, see skills and attitudes as the primary goals of the course for children. Teachers especially often stressed attitudes toward self and interest in learning as important outcomes.

Both teachers and regional directors agree content is important mainly as a mechanism for instilling attitudes or skills.

Only program often cited as an existing social studies curriculum that could do the same thing for children that MACOS does, was Taba. Complaint about Taba was that it has method but no pre-packaged materials.

Many teachers felt they could impart same attitudes and skills through other materials or courses. Others felt this ability was a result of exposure to MACOS training. Courses most often mentioned as possible vehicles were language arts, science, reading.

Teachers

Many administrators see MACOS as a good teacher training device, as do regional directors. See it as helping teachers develop child-centered teaching skills. In many instances teachers selected or volunteering for training have been, in principals' or administrators' opinion, among the better teachers in a school.

Teachers felt MACOS had affected their teaching style; made them more open to children's inputs. Often younger teachers felt they had been exposed to MACOS pedagogy in college methods

course and that their teaching style was less affected than did teachers who had been teaching for longer periods of time.

Regional directors and administrators said MACOS spread of effect was evidenced by teachers' change of role in classroom (less teacher domination). Teachers expressed same sentiment in different terms; "Am more open to students' questions"; "can tolerate 'good noise'"; "encourage debate rather than seek 'right' answers."

Often teachers would suggest that they try to tie MACOS in with the total classroom, approaching other subjects with discovery and student-centered methods, or using MACOS topics as starting points.

Some evidence of spread within a school; non-MACOS teachers would notice what was happening in a MACOS class, would express great interest in getting the course. Evidence of this is scarce, however.

B. Value of the course with different ethnic groups

Very little was found on this point; an isolated few white teachers felt materials have different values for black children. Black teachers felt materials had universal values. Asked if it could be possible that they don't know how white teachers use the materials, they said 'no.' They were trained with white teachers, often had middle class children to work with during training institute.

C. Values of the course across grade levels

Most teachers use materials on 5th or 6th grade level and feel it is well suited to those levels. Those teachers using MACOS on 4th grade level are split in their opinions. One group feels materials are too sophisticated for 4th graders. Others feel materials can be used with 4th graders successfully but not in as deep a fashion as with 5th and 6th. (One teacher said he would like to use materials with children in 4th grade, then again in 6th. Felt there would be something new to gain at each level.)

Some teachers using program on junior high school level with success. Feel it's a good introduction to anthropology. Also used as low as 2nd grade with success. One problem--where MACOS replaces traditional 6th grade social studies some teachers feel children are penalized because they will be tested at the end of the year on traditional social studies content for junior high placement. In a few cases, teachers tried to teach both MACOS and the regularly prescribed curriculum.

D. Teacher Pre-service preparation

Regional directors have conducted some pre-service training sessions. Either university classes or summer sessions, but evidence of their value is scanty. No instances of pre-service teachers later becoming MACOS classroom teachers evaluated that we know of.

E. Institutional change

Mixed opinion on this from regional directors. Some feel institutional change is a legitimate function of the course. Others feel no major change will come through only one course.

Those who feel that institutional change is a legitimate function of the course see it affecting:

school organization, e.g., team teaching
evaluation procedures, e.g., doing away with report cards
classroom norms, e.g., establish different norm which will ultimately establish other patterns of school behavior
teacher/principal roles, e.g., some principals have become more involved with students, teachers have developed new expertise and have been used as resources in other schools and systems.

One regional director felt the program could become a vehicle for upward mobility for individuals, but that it moves them up and out of the system rather than changing the system.

Teachers provided some evidence and information in the areas of institutional change. A few say they have abandoned A-F grading system. Several say that group planning or teaming has resulted from a need to cooperatively work out the schedules for films, etc. In some areas teachers are selecting other social studies curricula based on the adaptability to MACOS concepts and techniques.

In goals section, many teachers mention wanting to have or to develop independent study, non-teacher directed (new teacher role).

In curriculum selection, one administrator feels MACOS may be the answer to bilingual, bi-cultural problems; feels MACOS is culture free.

V. Evaluation

A. Many teachers said they didn't use traditional "tests" to evaluate students' progress with MACOS. However, rarely had they given up traditional grading for other courses as a result of MACOS. Often teachers would express concern over the question of measuring children's progress in any objective fashion; most felt evaluation had to be subjective and based on the teacher's knowledge of her children. No teachers could suggest any test that would adequately measure effects or impact of MACOS.

B. Administrators most often took the position that teachers' opinions or subjective evaluations of the course' merit was their yardstick, along with observations of children, reports from specialists, parents, and others, and their own judgment of the curriculum.

C. In only one instance that we found was MACOS being compared in a systematic way with other innovative social studies curricula.

D. No evaluation being done regionally with the exception of ERIE. Lack of funds and staff seemed to be the primary reasons.

E. When asked about evaluation for continued use of MACOS, administrators again were relying on teachers' opinions.

In the next section, illustrative comments and more detailed information concerning certain issues are given.

IV

FURTHER EXAMPLES AND TRENDS FROM INTERVIEWS

What do teachers perceive as goals of MACOS?

One of our interests was what teachers and others see as the goals of the MACOS program. We would typically ask teachers what they saw as the goals for their students and received, as one might expect, a great variety of answers. Some teachers felt that the goals are primarily cognitive; learning to think, to make generalizations, draw inferences, analyze, make comparisons, etc. Others used these process terms; learning how to formulate questions from data, how to verify, to collect and organize data, to observe. Some referred to social skills, such as learning how to listen to others in groups, how to participate and express themselves in groups, how to take part in discussions. Some teachers stressed feelings and attitudes as first concerns, especially with their slow learners or poor readers.

Exerpts from some of the interviews may help illustrate what we are talking about. A sixth grade teacher responded, "My goals are to try to get the students to look at themselves; build tolerance for people different from themselves; to use inductive reasoning; to teach them to compare different things--man vs. animal, for example. The process is more important than the content."

A teacher of fourth and fifth grades: "Concepts. What makes man human; the humanistic approach; have children deal with these higher order cognitive skills such as discussion, handling data, generalizing, selection of pertinent data. [To view] content simply as a supportive concept."

A sixth grade teacher: " [The course should help students to] think on their own without looking to teacher for direction or reinforcement. Children in not-so-structured classes are more used to group work than other

children, with ability to take charge. Social relations is a super-goal."

A group of sixth grade teachers: [The goals are to]"bring out the the uniqueness of man; appreciation of self, the human race; attitudes. Content is simply a means to an end. Students should think about self as as a human being that does have a place in society; there should be an appreciation of values and beliefs. Get them to think, to reason as they have never done before; develop understanding of behavior, provide interaction so they can observe changes in their own behavior."

A teacher of fifth and sixth grades: "The children learn to make generalizations, draw inferences, develop ability to participate in groups, learn to feel comfortable with questions for which there are no right or wrong answers."

Two teachers of non-graded children aged ten to thirteen: "Learning discussion techniques, to differentiate between fact and opinion; the course arouses curiosity, induces follow-up to activities. They learn to deal with values; no right or wrong. Build up critical thinking, give children opportunity to think; learn to defend opinions. [Learn about] human relations; everyone has differences, similarities. Teaches openness, respect for others."

A teacher of fifth and sixth grades: "In reading and math we group by ability; in MACOS we don't have to. Slow learners have discovered they can think too. Most value for these children. No right or wrong answer; every child has ideas. Most important goal is ability to think on one's own; to learn to listen; to respect opinions, not to say just anything. The course teaches children to think."

Many teachers commented on the value that MACOS had for different children. For example, they noted that there was something for everyone, and everyone could contribute. They felt that the course was particularly good for "poor readers" or "slow learners" in a number of ways. Because of the multi-media design of the course, children are not dependent on reading as their main or only source of information. Youngsters having difficulty with reading and writing could contribute by offering their thoughts and opinions. Many teachers felt children who had thought of themselves as unable to participate found that they could have "their day in court," as one teacher put it. "They have their chance to shine."

What seemed very clear was that different teachers saw different values in, or had different goals, for the course. In some cases, they believed that the priorities they had for the course were directly associated with the related needs of the children they taught. In other cases, teachers did not feel there was a direct relationship but that the goals served the needs of most children everywhere. Some teachers expressed the belief that the value of the course was that it enabled students to probe subjects about which they were relatively closed-minded. For example, one teacher said she had found MACOS offered a way of opening up discussion of the racial and poverty issues in the community; that is, she had found the course served as a vehicle for talk about people from ethnic and socio-economic groups other than those the children belonged to. She felt it was the only opportunity youngsters had to hear about and discuss people outside their own neighborhood and community.

The course was reported to be a vehicle for opening up discussion of various value issues otherwise difficult to deal with. One teacher commented, "We're always talking about education as dealing with values, but we never 'do' it." An elementary curriculum director said she had sat in

classes where children had been discussing aggression and hostility in themselves, feelings she had not thought about until she was over twenty, as she put it.

How are teachers evaluating the progress of students in the course?

One of the greatest concerns teachers expressed was how to assess progress. Those we talked with felt that grades were inappropriate. In some instances the teachers had refused to give grades. One group, under pressure to provide some kind of feedback to parents, took it upon themselves to have parent-teacher conferences to describe what the children were doing and how progressing. Some teachers have instituted a system of written reports to parents; others have worked out a satisfactory-unsatisfactory form of evaluation. Still others have had to assign grades at the schools' insistence; those we talked to were not at all happy about having to give grades.

Teachers gave us a variety of statements describing what they used to assess progress. Some periodically used a brief vocabulary or knowledge test of their own devising. In a few instances the teacher used the test developed at EDC. Some teachers have followed EDC's suggestion for interviewing children, or otherwise talking with them about their progress. Many said they based their judgment on what they saw the child doing, and on the extent to which the child participated. The kinds of questions he raised, his role in group discussions, the enthusiasm and interest he evidenced were all considered. Many teachers we talked with felt that tests of any kind were inappropriate or irrelevant. Some used self-made tests periodically as checks on progress; here a few noted that they used them mostly to have something to show to school or parents if need be.

Evaluation poses a dilemma for teachers. On the one hand, they know when children are interested, enthusiastic, and actively involved, and for many teachers this involvement in itself seemed to be a major accomplishment. On the other hand, the problem of how to communicate progress disturbs some teachers. Furthermore, the way to improve their handling of the course or to improve the course for certain children was unclear. One teacher commented that training for the course should include learning how to lead small group discussion. She saw her lack of training in such discussions a limiting factor when she first started MACOS.

When pressed, teachers would describe cases in which they felt the course had not been successful with certain students. One estimated that of thirty students around twenty-five would eventually become involved in the course. Some described youngsters who never really became interested. We were told about a class which remained uninvolved so that the teacher finally abandoned the course. In short, we obtained some limited coroborating evidence of the ranges of preferences and reactions of children which were described at length by Hanley, Whitla, and others in their evaluation of MACOS.

What do administrators, curriculum directors, and supervisors look for?

As with teachers, there are different points of view and different expectations. It is probably not unfair to say that many administrators see the course as a vehicle for changing teachers' attitudes and behavior in the classroom. This is not at all to suggest that they are not concerned with values for children, too. When asked what they would look at in evaluating MACOS, many administrators and curriculum directors said they would be guided by reports from teachers, and by observations

they made of children. A dilemma for administrators and others is that they are dealing with a one-year course. Problems of what to do after MACOS, and how to restructure the whole elementary social studies curriculum are of concern to various administrators. No one we talked with had any formal plans for evaluation.

It is worth noting that some supervisory and administrative personnel felt specifically that one of the desirable features of the curriculum was that materials are there. In their opinion, there are many aids to the teacher to help develop the pedagogical approach encouraged by the course; also, one does not have to spend an inordinate amount of time assembling additional materials. By the same token however, some of the frequently recurring problems mentioned by teachers have been lack of films, films not being available when needed, lack of a projector, late delivery of booklets, etc.

Can anyone teach MACOS?

Most people we talked to thought probably not, although the reasons given varied considerably. One center director, for example, felt that strictly reality-oriented teachers, to use his terms, probably could not and should not try to teach MACOS. By "reality-oriented" he meant having a need to deal mainly with facts and correct answers. The director also noted that strictly reality-oriented children probably should not take MACOS.

As one way of approaching the question of who could or should teach, and who he should teach, we asked a number of people how the course might fail. The most common answer was, "if it is taught as a traditional course by lecture, with emphasis on facts, and with the teacher engaging in single response questions and answers." In connection with this question

we hazard the guess, based on discussion with center directors and others, that so far the majority of teachers who have been teaching MACOS have been a fairly select group, either handpicked by the school system, or volunteers.

We also hazard the guess that this question will become increasingly important as the dissemination of the curriculum increases. We found one case, for instance, in which the board of education in a district adopted MACOS for all fifth grade classes next year. A principal in this district was puzzled by a small group of teachers who had made it clear that they did not want to teach MACOS. One reason they gave was that they had seen last year's MACOS teacher spending a considerable amount of time borrowing and returning films, and otherwise engaging in what seemed to be a good deal of extra and unnecessary work. We are not trying to say that these teachers will not want to teach MACOS next year; at this point in time they do not look forward to it. Even in groups that have taught MACOS, there are teachers who are not entirely enthusiastic about it. We were told of one who had not cared for the course at all, had been very dissatisfied with it, did not like its orientation and structure, and had felt very negatively about it by the end of the year. We do not know how the teacher's class felt about it.

The general point is that behind the patent success and acceptance of MACOS there are some real uncertainties relevant to adopting and implementing it. There are few guidelines to fall back on in working through the implicit decisions. It is not just a question of whether teachers are traditionalists or not; a number we talked with openly and emphatically declared themselves traditionalists previous to teaching MACOS, and now are entirely differently oriented. Some commented bluntly that MACOS had shown them how poor their teaching approach had been. In observing classes of teachers who describe themselves as previously very traditional, it seemed

clear to some that whatever "traditional" meant before, it does not apply now. It is not just a question of simplistic classifications; much depends, it appeared to us, on the training that a teacher receives in the use of the course, both before and during presentation of it, as well as on the teacher's needs and orientations.

Context Considerations

MACOS is a one-year program, and as such raises a number of questions about how it fits into various school contexts, programmatically and institutionally. With many of the people we interviewed, we tried to explore dimensions of the relationship between MACOS and the contexts in which it is being used. We asked teachers, for example, "Do you think the abilities and attitudes that children develop with MACOS will continue as the children move into the higher grades?" There were different responses. Some teachers felt pessimistic, their general position being that it all depends on what teachers of higher grades do. They thought that if upper grade teachers were traditional and approached social studies as a lecture course, then there would not be enough carry-over from MACOS. Some teachers hoped that the children would put pressure on subsequent teachers to continue with the kinds of instruction and learning that had been used with MACOS. Teachers of children who took the course last year are reported by other teachers to be impressed by the children's immediate interest in discussion and free contribution of opinions and ideas. The MACOS children, according to these teachers, seemed more ready to react positively to new subject matter in social studies.

Most teachers, however, who have taught MACOS before did not know what had happened with respect to children in subsequent years, although nearly everyone expressed great interest in knowing the outcome, and a need

for feedback. Lack of feedback was particularly noticeable when MACOS was taught at the sixth grade level, with children moving on to another school for seventh grade. In departmentalized teaching situations it is our impression that there may not be much cross-feed and feedback.

There were teachers who felt that MACOS builds a good foundation and potentially develops lasting attitudes. These teachers believed there would be carry-over of effects almost regardless of what happened subsequently. In one case, teachers and principal had worked on a follow-on program to MACOS to be implemented in sixth grade. The follow-on was built around further development of the five humanizing forces emphasized in MACOS.

We also asked teachers "Are there other subjects or curricula in this school at this grade level that may develop the same or similar knowledge, skills, or attitudes in children as MACOS?" Here again there was a variety of answers that seemed to depend, among other things, on the teaching situation and the school organization. Generally, we found that where teachers saw some continuities they were with science, reading and English language arts, and occasionally mathematics, especially when manipulative materials or enactive approaches are used. Where classrooms were self-contained, teachers were very likely to say that they were using more child centered approaches to teaching in various subject areas, applying where and as they could techniques they attributed to MACOS. In a few schools we visited where there were highly flexible arrangements such as team teaching, and multi-aged , multi-graded organization of children, teachers were likely to see most of the whole school program as consistent with and supportive of MACOS instruction. Teachers in self-contained classrooms were apt to regard themselves as very different from the rest of the school staff and program ("this is a very traditional, conventional

school"), and to regard little else as related to or consistent with MACOS.

With respect to social studies taught either prior to or concurrent with MACOS, the typical response was that MACOS was entirely different from the rest of the social studies program. Many teachers contrasted it strongly with what they regarded as traditional, uninspired social studies curricula. We encountered a few cases in which other, new social studies projects had been tried, such as the Taba social studies curriculum, and teachers seemed inclined to regard MACOS as similar but better in the sense that content and material were available to support methodology.

We asked teachers whether they had found any aspects of the MACOS training or curriculum had affected their teaching in other subjects. For departmentalized social studies teachers the question was not applicable. For others, there had been a spread of effect as described above, or else it was too early to tell. A problem for administrators and curriculum directors particularly arises over the necessities of rescheduling social studies that would normally have been taught during fifth grade, for example. This was not just a matter of subject area re-adjustment, such as rescheduling American History. In some cases the question arose about where to pick up the teaching of such things as library skills, graphing skills, and the like that were viewed as integral components of prior social studies programs. In this connection, one of the regional center directors noted that a key point one should remember in going before a board of education with a curriculum is "to think through and be prepared to discuss the implications, and how it will affect the rest of the program."

Views and experiences of the regional center directors

We explored a number of aspects of dissemination, distribution, use, support, and evaluation of MACOS with regional center directors. We won't try to describe in detail all of what was learned, but certain points are worth noting here. . . .

One question we were interested in was what standards of usage or procedure had been required or encouraged by regional center directors. There are various aspects of this question, and responses varied somewhat from center to center. Here is a summary:

"EDC requires five teachers for each film, but the practicality of the requirement depends on distribution and size of classes, and commitment of the central administration. There is a need for the curriculum coordinator to be disseminator and be directly concerned with decisions in dissemination, training, and quality control. The curriculum coordinator's authority is needed, and should be established ahead of time."

"The amount of instruction time required is forty-five minutes a day each week, or one hour four days a week. Haven't had other design standards; the course itself demands certain arrangements."

"Might be desirable to require or encourage standards of procedure, but whether that is practical depends on bureaucracy involved in the county, and the number of teachers who want to try MACOS. Unless the regional director works with a school system where he can be selective, he takes what he can get."

"We encourage a commitment from the district to use the MACOS materials to implement the program. We encourage a team, with at least one administrator, to come from each school district to the summer institute.

The team should be strong teachers with good leadership potential and good relationship with their peers."

"I feel it's the job of the regional director to visit the school principals and administrators to help create the support necessary."

"We demand attendance of every teacher at a workshop if she's teaching the course for the first time. Workshop models and format vary. We encourage non-MACOS teachers to attend workshops for credit. There should be entrance criteria for teachers to get into MACOS, but why exclude resistant, traditional teachers?"

Our observations within and among regions did not lead to any unequivocal impression of relationships between dissemination or training and performance. There were indications, however, that selection, training, and utilization processes and standards are important, can have effects, and could be further investigated.

Our general impression, based on this brief study, is that the MACOS curriculum is an extraordinary achievement on a number of counts, but that knowledge about its use is presently characterized by a large degree of indeterminacy or uncertainty and only partial explication of its educational properties and effects.

Overall, MACOS appears to have tapped several major trends or thrusts taking place in the social studies domain in particular, and in educational philosophy in general. Within the social studies program at the elementary level, there is a growing thrust toward seeing social studies as a vehicle for teaching inquiry skills pertinent to later effective citizenship. The approach described by Massialas and Cox would appear to be an example of

this thrust.[1] There is also a strong trend toward development of emotional and imaginational growth and competence in the service of educational objectives, a position exemplified by Richard Jones.[2] Underlying other interests in social studies is the concern with the continued imparting of subject area knowledge in a traditional sense. Yet another set of trends seems to concern a relationship between social studies and other areas, such as science. Here there is the position taken by Gagné[3] that process skills applicable in physical science are readily generalizable to other areas of science, including presumably the social sciences; however, Jones proposes the hypothesis that there are essential differences in the focus of the physical sciences and the social sciences from the point of view of the learner, with consequent differences in the nature of success and failure in the two areas.

---

1 Massialas, Byron G., and Cox, Benjamin C. Inquiry in Social Studies. New York: McGraw Hill, 1966.

2 Jones, Richard M. Fantasy and Feeling in Education. New York: New York University Press, 1968.

3 Gagné, Robert M. Psychological Issues in Science--A Process Approach. In American Association for the Advancement of Science, The Psychological Bases of Science--A Process Approach. 1965, p. 1-8.

V

SOME CRITICAL FEATURES OF THE MACOS CURRICULUM

These appear to be some of the critical features of MACOS that were expressed directly or indirectly by various people in the field.

1. The absence of a textbook as the core material
2. The central role of films
3. The scope, depth, quality, and applicability of available materials
4. The pedagogical and educational orientations of the course
5. The training component of past dissemination activities
6. The conceptual structure of the curriculum; multiple disciplines, spiral curriculum design
7. The multi-media design
8. The autonomous, self-contained package characteristic
9. The implicit and explicit value orientations
10. The lack of explicit outcome criteria and evaluation uncertainties
11. Lack of right-wrong answers; teachers no longer experts
12. Appeal to adult frame of reference and intellectual curiosity
13. Group orientation
14. Lack of closure--divergence--and the problem of subsequent continuity
15. Focus on humanity, inter-cultural relationships, humanness

VI

SOME CONSIDERATIONS CONCERNING FURTHER EVALUATION OF MACOS

After considerable thought, examination of a number of aspects of the present state of affairs with MACOS, and brief exploration of trends and issues that seem to be taking shape in social studies education in the elementary grades, we offer the following classification of possible courses of action for the National Science Foundation to consider.

1. What are the performance characteristics of MACOS? (i.e., how effective is it? what are its various outcomes?)

    a. with students
    b. with teachers
    c. with schools

2. What are the performance relationships of MACOS to other curricula?

    a. social studies
    b. science
    c. other

3. What are relationships of dissemination, training, and support systems to MACOS performance?

4. What are the relationships of MACOS to institutional change?

5. What are the implications of MACOS for other curriculum development projects or approaches?

6. What are the relationships of different evaluation strategies to MACOS performance?

Each of these questions will be discussed separately.

1. Performance characteristics of MACOS (how effective is it?)

There are seven major goals stated for MACOS.[1] Of these, six refer to students as the target group, one refers to teachers. The goals are unspecified in the sense that there are no criteria provided for deciding whether goals are accomplished, approached, or missed entirely. Much is left to the teacher to judge. The Strategies for Evaluation now available should alleviate some of the uncertainties and perplexities so frequently expressed by teachers. The manual contains the short content questionnaires (tests) that were used in the formative evaluation studies of Hanley, et. al., as well as classroom checklists, interview guidelines, guidelines for use of creative formats in assessing progress, and checklists for classroom observations. There are, however, no declared priorities of goals, although it seems clear enough that goals of methods and attitudes are, generally, of paramount priority.

It seems to us that the crucial characteristic of MACOS from the point of view of performance evaluation is its duration of one year. It is a package that assumes no prior preparation of pupils and does not necessarily directly connect with subsequent curriculum subjects or units. The skills and orientations that children are encouraged to develop in MACOS are presumably relatable to prior and subsequent skills and orientations. We say presumably since it is by no means clear to us that the procedures and criteria are specifiable for defining, evoking, and identifying inquiry skills in some ordinal fashion; or in some generalizable fashion from one subject area to another; or from one time period to another. As Cole and

---

[1] Education Development Center. Man: A Course of Study Evaluation Strategies. Cambridge, Mass., 1970, p. 12.

Seferian put it, "one man's inference is another man's classification."[2]

There is a need for careful pre-course observation and measurement, as well as for follow-up well beyond the end of the course. The pre-course observation is more than a matter of pre-testing for knowledge, intelligence, reading achievement, etc. It should include establishing performance trends in social learning behavior, inquiry skills, and attitudes or interests of children prior to their taking MACOS. The follow-up assumes especial importance under the condition in which children transfer to another school upon completion of the course, as happens frequently when it is given in the sixth grade. The presumptions are 1) that the course may have different performance characteristics with respect to any or all of its avowed goals with children of different previous habits and expectations;[3] and 2) that the course will have differential extended consequences according to the different fates of children beyond the grade in which it is given.[4]

The time-span of the course (and its implications) is, in our opinion the single invariant property of MACOS bearing on continuing evaluation.[5]

---

[2] Cole, Henry P. and Seferian, Albert. Analysis of process curricula. In Andreas, Burton G., et.al., Research in Process Curricula. Eastern Regional Institute for Education, Syracuse, New York. March, 1970.

[3] The Hanley report consistently documented differences in performance between inner city and suburban children. It called attention to different entry characteristics of children in these categories. These are gross categories. No analysis of variations within categories was made.

[4] In each case, extended pre- and post-course observation would bear not only on questions of differential performance characteristics, but also on questions of what schools might do before and after the course to maximize its usefulness.

[5] There are, to be sure, other invariant characteristics of the package. A conspicuous one is the absence of a textbook and the centrality of films. The availability of the films and other materials to teachers as needed may have consequences for morale and performance. Another invariant, as already noted, is the absence of specific objectives and goals.

There are other characteristics relating to continuing evaluation that are not immutable or invariant. One, of central importance, is the unfamiliarity of many teachers with the content of the course, and to a lesser extent, with the pedagogical requirements and roles. The evidence to date is that students are having difficulty mastering concepts, and that the methods of inquiry are neither extending beyond the boundaries of the course nor enduring. This may be in part a consequence of developmental limitations of children. Until it can be shown, however, that teachers have mastered the course conceptually as well as pedagogically, the newness of the structure of the discipline to teachers has to be considered an important characteristic in continuing evaluation.[6] The implication here is that the growth qualities of teachers need to be examined. This is not as straightforward as might appear. Departmentalized teachers in fact may teach each unit four to five times as frequently as self-contained classroom teachers during the same calendar period. Most of the departmentalized teachers we talked with felt that each succeeding class during the day was different per se and for the teacher. One or two said that they sometimes felt like a tape recorder at the end of the day. Teachers in team teaching situations of one form or another may be more like departmentalized or more like self-contained teachers, depending on the arrangement. The potential difference with team teachers, however, in speed of mastery of the course, is the opportunity to engage in more continuous examination of its thrusts and concepts through exchange of observations, ideas, information, and opinions.

---

[6] As Goodlad has noted, elementary school teachers are typically trained as generalists. Goodlad, John I. Direction and redirection for curriculum change. In Leeper, Robert R., ed. Curriculum Change: Directions and Process. ASCD, NEA, 1966, pp. 1-14.

Another variable characteristic of the course that has, in our opinion, important implications for continuing evaluation of performance is its pedagogical and content linkage with other segments of the school day from the point of view of students. It is not unreasonable to hypothesize that educational situations that reinforce behaviors, attitudes, and skills reinforced in MACOS should enhance at least the generalizable behavioral and cognitive outcomes of MACOS. This continuity or discontinuity of opportunity, utilization, and reinforcement patterns is not necessarily correlated with the organization of instruction (i.e., self-contained; departmentalized; team teaching, etc.).

A third variable that seems especially important is the large degree of internal flexibility of teaching arrangements and processes inherent in the course. The course itself is highly structured with respect to the sequencing of units. The options for proceeding, the pacing, the formation of groups of students, the opportunities for exploiting topics and issues of interest that may arise, and so on, all are sources of variability that potentially bear on performance of the course. It is this flexibility obviously that makes it difficult to think of MACOS as a treatment. This characteristic, in our opinion, poses methodological problems for evaluation compared to which most others seem pale.

A final variable is the focus on the group as the operating unit of the course. This is the design consequence of Bruner's concept of the function of the reciprocally operating group in learning. An implication of this focus is that, while the individual student is the carrier of the results of education, the group may be the more appropriate unit for observing and evaluating some of the results of the course.

The preceding points suggest that further evaluation (of the extent to which the course achieves its goals, given that they can be operationally defined), should take into account the prior context of students' education, the concurrent context, the subsequent contexts, the teachers' mastery of the course, and the internal organization and conduct of the course, as well as the developmental level and abilities of students. Further evaluation should also use a variety of measures and observation procedures.

If it were possible to show reliable and operationally meaningful differences in performance (effectiveness) with respect to any or all of the goals in relation to various operational conditions, on what decisions would such information bear? Essentially, continuing evaluation would support decisions at local school levels by providing baseline information about performance under various conditions of input and implementation. It would also relate to some decisions about further curriculum design and development.

2. What are the relationships of MACOS with other curricula in social studies, and in other subjects or disciplines?

In the preceding discussion, the MACOS package was treated as a device as such. The questions arising from this concept involve what the device does (or enables people to do), under varying conditions, and under varying time periods. A second point of view is a consideration of MACOS as a member of a class, as an exemplar of a set. As such, it can be examined for its class implications, and it can be examined comparatively in relation to other members of its class.[7]

Viewed from this perspective, MACOS has certain characteristics that seem especially important for further evaluation. One is the focus on mechanisms for survival, or more broadly, on structures and functions vis á vis environments. A second, which is a corollary of the first, is the focus both on social and individual behavior and their relationships to motives and feelings. Two others of particular significance are the multi-media design (and the realistic, non-expository nature of the films) and the extensive array of material in support of content and teaching process (including lesson planning and preparation). A final characteristic for comparative evaluation is the spiral design of the curriculum.

The implication of the focus on mechanisms for survival is that the course has a basic value orientation, or at least a criterion for making judgments; namely, survival. In this respect it is not simply descriptive. Therefore, if it were to be compared with other social studies courses, for example, on whatever measures, it would be appropriate to consider specifically the role of the underlying value orientation (or the

---

[7] The various classes of which MACOS is a member include process curricula; inter-disciplinary curricula; student-centered curricula; multi-media curricula; social studies curricula, etc.

presence or absence of one).

The implications of the focus on behavior are for the motivating efficacy of the course, as well as for educational outcomes. The multi-media design, diverse learning materials and spiral design have implications for measures of efficiency as well as effectiveness. It is possible that, other things equal, the multi-media design has the effect of broadening the range of students who achieve particular levels of vocabulary, conceptual mastery, interest, etc. Similarly, the availability of materials may in fact have effects not only on learning time for teachers, but also on general acceptance by teachers.

These considerations suggest possible directions or forms of comparative evaluations. In addition to comparisons of performance in cognitive, affective, and behavioral domains with students, and in similar domains for teachers as teachers, it would also be appropriate to compare MACOS with other class exemplars in terms of scope of effectiveness, installation efficiency, and implementation efficiency. Comparative evaluations would support or relate to adoption, deployment, and support decisions at local and state levels. They would have greater bearing in the long run, however, on curriculum design and development decisions, we believe, and possibly on decisions concerning teacher preparation and in-service training.

3. What is the relationship of MACOS' performance to dissemination, training, and support?

The cluster of Regional Centers established for MACOS has been extremely important in the success of the curriculum to date. While there has been substantial variation in approach among regions (within the general limits of the model set up by EDC), there is ample evidence of the values of the roles played by the Regional Centers, and of the implications of variations in roles. Yet there are few data readily available bearing on evaluation of roles and methods. Only one center has systematically collected and analyzed information about the use of the course by teachers within its region. There has been wide variation in record-keeping among centers, and no requirement and funding for basic data recording related to parameters of dissemination, training, and support.

Do dissemination, training, and support affect performance? Are there optimum approaches? In what ways do they depend on characteristics of the receiving and sustaining environments? In what ways do they depend on the curriculum or materials?

Obviously, dissemination affects performance in a trivial sense: no dissemination, no performance. Beyond that, there are a host of questions about processes of dissemination, adoption, and support that can eventually be related to performance.[8] The need to train teachers to use the course seems self-evident. But how effective are different training models and methods, with whom? There is evidence that some teachers who go to the MACOS summer institutes to become teachers of teachers during the school year may not, in fact, follow through with the expected seminars. Is that because there is inadequate incentive? Inadequate planning and prepara-

---

[8] See, for example, Carlson, Richard O. Adoption of Educational Innovations. Center for the Advanced Study of Educational Administration, University of Oregon, Eugene, Oregon, 1965.

tion? What is the consequence for the performance of the curriculum with teachers as well as with children? What are the consequences for performance of different adoption processes or modes? So far the course has been very successful largely with teachers who have volunteered to learn and teach it.

What constitutes support? One teacher with whom we talked specifically recommended that the regional center should be much more active in supplying information and literature about MACOS utilization, techniques and developments. The MACOS newsletter established by ERIE for its regional center is one excellent form of such support. In another case, we talked with a district director of social studies who was working with teachers using MACOS on the problems of evaluating discussions by students over time. That is a form of support for the course. One principal expressed concern over the lack of publicity and feedback concerning the use of the course nationally. Is a form of support needed here and would it affect performance? These examples are intended to convey a range of possibilities beyond the obvious logistics and supply forms of support.

Some of the characteristics of MACOS that may be especially related, or possibly sensitive, to dissemination, training and support strategies and methods include:

a. the clash of the course with established curricula and with value systems

b. the interplay between the highly organized structure of the course and the reinforcement or encouragement of divergence among children in its implementation

c. the emphasis on and opportunities for engagement of emotion and cognition in children toward educational goals

d. the shifting or multiple requirements of the roles of the teacher

e. the unpredictable interplay of needs and expectations of children with the orientations and requirements of the course

Would evaluation of dissemination, training, and support systems or models be feasible at this time? Our opinion is that it would be, perhaps with special arrangements with the publisher, Curriculum Development Associates. If the focus is on generalizable variables and models, the value of such study could be great in the long run.

Evaluation of various adoption, training, and support processes and models would relate directly to adoption and implementation decisions. It would also relate to decisions about the requirements of overall development, distribution, and support systems in future curriculum production efforts.

4. Relationships of MACOS to institutional change

The curriculum can and does have affects on schools and communities. In the case of MACOS these have covered a range of reactions, from highly positive ("for the first time my youngster is talking at home about what he's doing in school") to highly negative (concerns about evolution, sex education, etc., occasionally emerge). Similarly, reactions of teachers, students, curriculum specialists and others have varied, as one might expect.

There is little evidence that we could detect of institutional change related to MACOS. There was evidence of changes within school buildings. The course has led some teachers to challenge the grading system, at least as applied to MACOS. Other examples of changes stimulated by the course are further curricular changes for earlier and later grades; changes in school-home or teacher-parent relations; changes in supply distribution activities; changes in relationships among teachers; changes in views of teachers, administrators, parents, etc., of what an educational course is

(where is the textbook? why don't you have homework in a book?); and so on. There are also potentially significant changes in the roles of the teacher and the student. Indeed, this is one of the goals of the course. But with the possible exception of this last point, these are modest changes, compared with the sweeping institutional changes and reforms so urgently and increasingly called for. For instance, the course may change the role of the teacher in the classroom, but it will not per se affect his status and role in the system. It will not affect the organization of educational activities, though it conceivably could contribute to redefinitions of educational goals. Nevertheless, on a local level some of the impacts of the curriculum may have ripple effects or synergistic effects over time.

One group at a regional center said it is expecting too much of one little social studies curriculum to ask that it have much institutional impact. They are almost certainly right. It is of interest, however, to individual schools, teachers, parents, and others to know more about the possible secondary and tertiary effects of curricula. Evaluation focussing on such impacts would provide information relating to decisions concerning strategies of dissemination, adoption and support locally, and at higher levels (state, regional, and national).

5. Implications of MACOS for other curriculum development projects

The Man: A Course of Study curriculum can be examined from yet another perspective. One may explore the implications of the MACOS project and its outcome for further curriculum design and development efforts. An evaluation of the curriculum from that point of view would encompass a much broader set of data than even a comprehensive appraisal of the course per se.

One knowledgeable director of instruction in a school district we visited asked how realistic it is to expect teachers to develop courses during the summer, considering the time, talents, and resources that went into MACOS. Does this country need a curriculum design and development machinery on the scale, say, of a space exploration program?

Goodlad has called attention to the direction that curriculum reform should take.[9] What kinds of funds and organizations are needed to move in the directions Goodlad has suggested? Would an analysis of the MACOS project and curriculum yield information leading to important decisions about the investment of resources and the creation of new alliances and inter-organizational arrangements? On a more mundane level, there are implications for curriculum design in the conjunction of content and teaching roles and methods. There are, obviously, implications in the multi-media form of the course, in the inclusion of games, in the forms and sequences of contrasts, in the completeness of materials specifically designed and developed around the conceptual structure of the course.

Evaluation of the project and the curriculum oriented toward this issue would bear to some extent on decisions involved in subsequent design. It would bear most directly on decisions about the allocation and organization of resources for curriculum change in the social studies and possibly other areas.

---

[9] Goodlad, John I., op. cit.

6. What are the relationships of alternative evaluation strategies to MACOS performance?

The purpose of the present study was to suggest alternatives for the National Science Foundation to consider with respect to further evaluation of MACOS. It is appropriate, however, to consider for a moment the subject of evaluation strategies as such. Let us shift the focus from the question of what strategies are appropriate or feasible to the question of the properties of evaluation models and approaches per se. Let us consider evaluation models as a subject for research and demonstration, with MACOS as the occasion or vehicle for the examination of different models.

There are several evaluation models currently receiving wide attention. While all evaluation presumably has to do with decision-making, different models seem applicable to different levels of organization, to have different operational implications, and to have different utilization characteristics. Stake's transactional model, for example, appears to be aimed at the immediate operational level--the classroom.[10] There are strong parallels between Stake's conception of evaluation and Jones' descriptions of experimental lessons in the development of MACOS.[11] Stufflebeam's

---

10 Stake, Robert E. The Countenance of Educational Evaluation. Teachers College Record, 68, April, 1967, 523-540. For two applications of Stake's general model, see Chapter 6 by Roland F. Payette and Benjamin C. Cox (New Dimensions in Evaluation of Social Studies Programs), and Chapter 7 by Irving Morrisett, W.W. Stevens, Jr., and Celeste P. Woodley (A Model for Analyzing Curriculum Materials and Classroom Transactions) in Dorothy McClure Fraser, ed., Social Studies Curriculum Development: Prospects and Problems. 39th Yearbook. National Council for the Social Studies, NEA, 1969.

11 Jones, Richard M. , op. cit.

CIPP model[12] focuses on a broader operational sequence and seems to lead to a wider range of decision-making levels. Provus' model[13] appears most applicable to decision-making at the level of the school district. This is not intended to be a comprehensive survey of evaluation models.[14] The point is that there are a number of alternative models with varying focuses and varying presumed operating characteristics. The MACOS curriculum, we suggest, could provide a vehicle for exercizing or examining several alternative evaluation models for the purpose of clarifying their utility and operating characteristics.

The curriculum is no longer in a developmental stage. Funding decisions concerning research and development of that particular curriculum are over. But the implementation, maintenance, and improvement decision-situations have really just begun. Furthermore, MACOS is sufficiently large and costly to have extensive educational and administrative consequences.

---

[12] Stufflebeam, Daniel L. Evaluation as Enlightenment for Decision-Making. Columbus, Ohio: Evaluation Center, The Ohio State University, 1968. See also: Stufflebeam, Daniel L. Toward A Science of Educational Evaluation. Educational Technology, July 30, 1968, 5-12; and Guba, Egon G., and Stufflebeam, Daniel L., Evaluation: The Process of Stimulating, Aiding and Abetting Insightful Action. Columbus, Ohio: Evaluation Center, The Ohio State University, 1968.

[13] Provus, Malcom. The Evaluation of Ongoing Programs in the Public School System. In Tyler, Ralph W., ed., Educational Evaluation: New Roles, New Means. National Society for the Study of Education. Chicago: University of Chicago Press, 1969.

[14] For others, e.g., see Evaluation For Administrative Action. Journal of Research and Development in Education, Vol. 3 (4), Summer, 1970, passim. Also, Stake, Robert E. Testing in the Evaluation of Curriculum Development. Review of Educational Research, Vol. XXXVIII, No. 1, February, 1968.

Yet it is sufficiently delimited to be readily observable. It is presently being used in a wide range of environments; there will in all likelihood be greatly increased adoption of it in the next few years.

The question posed here is, in what ways do various evaluation models or strategies affect MACOS performance? While an appropriate measure of effectiveness of evaluation processes is their influence on decisions (at whatever level), these decisions will relate to the utilization and effectiveness of the curriculum in one respect or another. Some evaluation models, especially those that are avowedly system analytically oriented, base their claims for utility in part on the power of feedback loops. An important question about such models is, what are the operating characteristics of such feedback loops? Feedback abounds. There is need to explicate planned or presumed feedback loops to improve the usefulness of educational evaluation.

VII

ALTERNATIVE STRATEGIES OR APPROACHES

In this final section we shall discuss specific approaches to further evaluation of MACOS. We shall offer at least some analysis of the advantages and limitations of each.

In the preceding section we discussed briefly six evaluation issues or categories of questions about MACOS. Here we shall be concerned with three major alternative approaches to further evaluation:

1. absolute evaluation--evaluation of the performance characteristics of the curriculum.

2. comparative evaluation of the curriculum

3. systemic evaluation--evaluation of the systems affected by or related to the curriculum (this incorporates points 3-6 in the preceding section).

These are not mutually exclusive choices, necessarily. Comparative evaluation subsumes absolute evaluation (or at least aspects of absolute evaluation). Either absolute or comparative approaches could include systemic evaluation. We think there are advantages, however, in considering them separately.

Absolute Evaluation of MACOS

MACOS was not designed according to a planned structure of specific behavioral objectives. It seems characteristic of curricula designed to facilitate open-ended teaching rather than goal-directed teaching.[1] It has seven goals that can be operationalized. There are 65 units, each capable of being evaluated. There are test items that have been used in the formative evaluation of the curriculum, and which can be used to measure certain goal achievements. The items are few in number, and it is not at all clear how they sample the universes to which they belong. They are highly specific to the content of the course. Norms are available for the Animals section, but not for the Netsilik section. The questions for that section are intended to be diagnostic tools for the teacher, not ranking devices. The evaluators, during MACOS' development, made extensive use of interviewing as a data gathering technique. They have provided teachers with constructive guidelines for using the same methods themselves.

Further appraisal of the curriculum, using as a minimum the methods and instruments employed in the formative evaluations, is obviously feasible, depending on the criteria of reliability and precision one sets. Our investigations have led us to suspect that further study designed to replicate Hanley's work will eventually show, perhaps more elaborately and in more readily interpretable form, that Hanley's results, both pro

---

[1] Cf. Lee S. Shulman and Evan R. Keislar, eds. Learning by Discovery. Chicago: Rand McNally, 1966. See especially the editors' summarizing comments about curriculum development, pgs. 186-190.

and con, are substantially reliable with respect to outcomes for students and some aspects of teaching style or technique and classroom management.

More instruments and items could be developed to yield more explicit and manipulatable measures of MACOS goal-related behaviors. These could and probably should include some performance tasks to be undertaken by groups or classes of children, as well as measures of individuals. Categories of possible outcomes of MACOS, even in a non-comparative evaluation, could and no doubt should be extended insofar as observation is concerned. Cronbach's suggested list serves as a useful starting point for such an extension.[2] It is not clear, however, that development of an evaluation model based on arrays and hierarchies of specific behavioral goals and goal criteria would be an undertaking worth doing for MACOS. The elegant model developed for Science--A Process Approach[3] applied to a curriculum with a theoretical base and content quite different from MACOS. There of course must be criteria for categorizing observations, and, in some cases, ordering degree or quality of complex behaviors such as hypothesizing, generalizing, listening to others, and so on. There are instruments

---

[2] Cronbach, Lee J. The logic of experiments on discovery. Ch. 5 in Shulman and Keislar, op. cit.

[3] AAAS Commission on Science Education. Science--A Process Approach: An Evaluation Model and Its Application. Second Report AAAS Miscellaneous Publication 68-4, 1968. Note incidentally, the discussion of risks, uncertainties, and potential values of this evaluation approach. (p.34-37)

available that enable observers to rate the occurrence of particular forms of behavior, or the operations presumed to intervene between stimulus and behavior. Criteria for classifying and rating responses can be applied to interviews and to the interpretation of paper and pencil test items. Measures of performance for MACOS with children can be structured into hierarchies of impact value in the following sense:

1. Acquisition: rate of occurrence and/or quality of particular dimensions or kinds of knowledge, skills, or attitudes (question posing, data gathering, categorizing, attitudes, orientations, concepts, models, etc.)
2. Lateral transfer or extension: amount and kind of application of concepts, ideas, orientations, methods, behavior, attitudes to concerns, situations, or topics other than those of the course during the child's involvement in the course
3. Retention beyond involvement in the course: recall of or savings time in relearning vocabulary and information, concepts, methods, etc.
4. Longitudinal transfer, extension, incorporation, or elaboration: amount and kind of application or continuity of use of content, concepts, methods, behavior, orientations, attitudes, etc. beyond involvement in the course

The stated MACOS goals refer only to the first level. The presumption, however, is that there will be other, more enduring payoffs such as those implied in the second and third levels and explicated in the fourth. In effect, acquisition, extension, and application (broadly) of knowledge, skills and/or attitudes, each conceived broadly also, may be viewed as a rough scaling of effectiveness that applies, potentially, to any course. It also applies to the effects or impact of the course on teachers.

Any evaluation study interested in the third and fourth levels

of effect or impact with respect either to teachers or students must be longitudinal. It must extend at least a year beyond the presentation of the course, and preferably several years.

MACOS is intended to be a process-oriented curriculum. Like most other terms, process is ambiguous in that it has a number of simultaneous referents. Children may learn hypothesis generating, exploring and testing processes, a family of behaviors referred to by various people as inquiry methods, research methodologies, etc. The means by which children learn or are taught such processes constitute another family of behaviors that generally include the teacher in one or more roles, functions, or capacities. What contingencies or functional relationships (in Stake's terms[4]) exist between these two families of processes is an empirical matter that could be examined by various evaluative as well as research studies. The nature of these relationships was, to be sure, one of the central relationships of the conference on Learning by Discovery.[5]

If evaluation of the performance characteristics of MACOS as such is to yield more information than is already available in the Hanley report, it should include consideration of at least the following:

General Design Requirements

1. Repeated observation over a period of at least three years of the same populations of students, teachers, schools and school systems. The need for follow-up with students is apparent from the Hanley report; the need for follow-up with teachers,

---

4 Stake, Robert E., op. cit.

5 Shulman and Keislar, op. cit.

schools, and systems was apparent from the informal information gathered by this study.

2. Explicit attention to different prior, parallel, and subsequent instructional programs, organizations of instruction and pedagogical orientations and styles to which children studying MACOS are exposed.

3. Explicit attention to the processes of adoption of the curriculum and the training and support of teachers.

4. Documentation of the variations in goals and expectations of teachers and school systems for the course. They vary, and the explication of congruences (again in Stake's terms) of expectations and outcomes, or of goals and examination content (in Scriven's terms[6]) is an important output of an absolute evaluation. There is a logical dilemma here, in that specific goal criteria are not established in operational terms for MACOS, other than those implicit in the content questionnaire. An appropriate strategy, under such a circumstance, would be to measure as many outcome variables as possible in order to examine differential relationships of goal expectancies and outcomes.

5. Inclusion of a variety of outcome measures. There should be measures of primary, secondary, and tertiary effects (again in Scriven's terms) as implied in the preceding points. But there should also, to the extent possible, be several forms of measurement of the same variables, especially with respect to performance of children. Methods should include classroom observation, paper and pencil tests, rating scales, interviews, and possibly the use of standardized group tasks or projects. Different methods will permit measurement of similar kinds of behavior and competency under different task conditions. The

---

[6] Scriven, Michael. The Methodology of Evaluation. In Perspectives of Curriculum Evaluation, AERA Monograph Series on Curriculum Evaluation, No. 1, Chicago: Rand McNally, 1967. Examination content would here have to refer to content of interviews, paper and pencil tests, attitude scales and other methods of observing and measuring course goal-related achievements and behavior.

conclusions one may draw about the acquisition of, say, inquiry skills, however defined, from responses to a paper and pencil test may not be the same as those drawn from the observation of the same groups of children responding to a structurally similar problem in an extended group problem solving situation, or in an interview.

Specific Design and Methodological Considerations

The preceding requirements delineated general sets of independent and dependent variables, as well as general observation periods. It is not our intention to suggest specific designs and methodologies within the framework of absolute evaluation. It is appropriate, however, to note some design and methodological considerations for National Science Foundation to include in choosing among subsequent courses of action, of which absolute evaluation is one.

1. Selection of schools, teachers and students for study.

We assume that the general question of interest in an absolute evaluation is, what does the course do with various kinds of students under various types and conditions of adoption and application? There are essentially two kinds of comparisons of interest; within subjects (students, teachers) over time,and among treatments (different organizations of instruction, types of teaching, types of schools, etc). It is unlikely in the extreme that there can be random assignment of schools, teachers or students to MACOS in anything cther than occasional local experiments. Furthermore, inasmuch as the curriculum is only just beginning to be disseminated and adopted on a competitive basis nationally, any sample of schools and school systems chosen for inclusion in a study at time t (say, this calendar year) has an indeterminate relationship to all schools or systems.

We suggest that an appropriate strategy is to try to define a current universe of using school systems, based on information available from the Regional Centers, EDC, and CDA, listing as well as possible the systems and schools meeting one or more conditions or levels of independent variables of interest. Then draw from this pool to the extent possible. This could be done within a single region, as presently constituted, with probably little practical loss of generality other than variations in

conditions and types of training for use of the course, and with considerable gain in access and reduction in cost. We hypothesize that the main loss would be in the face validity of results, based on the assumption that there is apt to be more variation within than between regions for most interesting variables related to the course.

Another strategy is simply to construct clusters of schools and classrooms meeting eligibility criteria on various conditional variables. This is in effect the approach used by the EDC evaluators.[7]

2. Size of the study

There is no simple way of designating the appropriate size of a study until one gets down to specific hypotheses and designs and parameter estimates. There are choices among units of analysis. For some hypotheses schools may be the appropriate units, for others, classrooms, and for others, students. When absolute evaluations are considered, it will be well to consider two additional factors in determining the size of the study:

a. follow-up with students in periods after their involvement in MACOS may be very costly if attempt is made to trace students associated with particular classrooms;

b. in a longitudinal study one can expect the attrition of teachers as well as students to be high; in some schools attrition may run on the order of 15 percent a year for teachers and two to three times that or more for students.

---

[7] It was also, as far as we can tell, the method of the ERIE researchers who are investigating process curriculum installation problems. See Andrulis, Richard S. Variables Affecting Installation, In Andreas, Burton G. et.al., Research Into Process Curricula. Eastern Regional Institute for Education, Syracuse, N.Y., March 1970.

We suggest that if absolute evaluations take the form of in-depth analyses, such as those implicit in Joyce's forthcoming report[8], something on the order of twenty to thirty teachers may be quite sufficient. Otherwise we would expect that numbers of teachers up to an order of magnitude greater than that will be needed.

Implementation Mechanisms

There are several mechanisms or vehicles for conducting a longitudinal evaluation. Each has its advantages and limitations. Mechanisms that National Science Foundation could consider include:

1. providing a grant to an institution or agency (presumably a university or university-affiliated center or agency) to conduct a specified study
2. providing redundant or supplementary grants to several such agencies either to replicate specific studies or to conduct complementary studies
3. setting up a MACOS evaluation project agency to conduct studies
4. extending the capabilities of one or more of the present regional centers to include a research and evaluation function
5. collaborating with other agencies (United States Office of Education; National Council of Social Studies, etc) to establish an evaluation project center, system or commission.

We shall not undertake a detailed analysis of these alternatives, which apply to other evaluation approaches to be discussed below. The following remarks are intended simply to indicate some pros and cons of various mechanisms.

---

8 Joyce, William W. MACOS: A Report from the Inner City. Interim report for Elementary Education Supplement, NCSS. (In press).

It is likely that the most economic mechanism would be # 1 or # 2. Either of these has the advantage of providing a centralized control of the entire process of specific studies. They have potential limitations in availability of technical people at irregular periods during the academic year. In general, however, either option represents a minimum investment and would thus seem to be appropriate vehicles if NSF envisions further evaluation as a terminal activity. If, however, provision of continuous and useful feedback to participating or cooperating schools is important to NSF, this is potentially the least effective mechanism unless special provisions are made.

Option # 3 has the advantages of the first two, plus the advantage of continuous concentration of resources and effort. It has the advantage of providing a focal point for feedback, and for the coordination of further development work, especially with respect to methods of measurement and instrumentation. Especially if NSF sees the ultimate aim of further evaluation as producing information for more curriculum development projects, an argument can be made that this option, or the fifth, are the most attractive mechanisms. The third option has the disadvantage, we think (but do not know), of greater investment costs than the first or second options. It also duplicates the capabilities, or potential capabilities, of already existing centers and agencies elsewhere. It could also lack objectivity.

The fourth option has a number of advantages. The regional centers have been the dissemination and training arms of the MACOS project. They have access to schools and to university resources. In the case of the ERIE center, there has been an active and impressive data gathering and feedback operation, even under the minimal funding available.

The disadvantage that comes most readily to mind is the essential conflict of interest in co-locating a formal evaluation function

with a dissemination and training function. Regional centers are, after all, in the business of disseminating MACOS and training teachers. If the essential aim of an absolute evaluation is to delineate the various achievements that the curriculum can have under various conditions, as Bruner has urged[9], then appending an evaluation function to the centers seems appropriate. Otherwise, and with no imputations concerning the integrity and objectivity of regional center staffs, the evaluation function should be kept separate from the operational functions, in our opinion.

A combination of options # 3 and # 4 could conceivably counteract the conflict of interest problem. Even if option #4 were selected, it would be necessary to have a coordinating and directing office or agency to provide uniform specifications and to assure comparability of results. The model here would be something like Bond and Dykstra's Cooperative Research Project in reading,[10] though probably smaller in scale.

The fifth option has a number of obvious advantages, especially if NSF sees itself moving more broadly into further curriculum development in the social studies, or into projects with a potential for institutional change. There are already those who are calling for large-scale, federally funded programs to manage research and development related to thinking skills and processes comparable to the Right to Read program.[11] Goodlad

---

9 Bruner, Jerome S. Some Elements of Discovery. Ch.VII in Shulman and Keislar, op. cit., p. 113.

10 Bond, Guy L., and Dykstra, Robert. Coordinating Center for First Grade Reading Instruction Programs. Final Report, Project No.X-001, Contract No. OE-5-10-264, Feb., 1967. ERIC ED-013-714.

11 Lundsteen, Sara W. Critical Listening and Thinking: A Recommended Goal for Future Research. Journal of Research and Development in Education. Vol.3 (1), Fall, 1969, 119-133.

recently called for broader, more coordinated efforts in curriculum improvement and reform.[12] Although this option is more germane to comparative or systemic evaluations than to absolute evaluations, it is not altogether inappropriate for consideration here. Its disadvantages are easy to imagine, if not verify. They include cumbersomeness, problems of inter-agency coordination, and, perhaps, from the NSF point of view, abrogation of competitive interests and advantages.

There are other conceivable mechanisms, to be sure. We think these are the principal ones likely to be considered.

Objectives and Payoffs

Stake ended his now classical paper on the countenance of educational evaluation with five questions that ought to be answered before undertaking a formal evaluation.[13] Four of the five apply to consideration of absolute evaluation of MACOS.

1. Is it to be primarily descriptive, judgmental or both?
2. Is it "to emphasize the antecedent conditions, the transactions, or the outcomes alone, or a combination of these, or their functional contingencies?"
3. Is it "to indicate the congruence between what is intended and what occurs?"
4. Is it "intended more to further the development of curricula or to help choose among available curricula?" [14]

We have already suggested that absolute evaluation of MACOS would seem to relate mainly to adoption and implementation decisions by schools, and to certain decisions concerning future curriculum design and development. Realistically, it would seem that the emphasis of an absolute

---

[12] Goodlad, John I., op. cit.

[13] Stake, Robert E. op. cit.

[14] Stake's other question--is this to be a comparative evaluation?--is answered in this discussion by definition.

evaluation of a curriculum that does not have specific performance standards or objectives should be descriptive. It should probably emphasize antecedent conditions and outcomes if it is to serve the purpose of contributing to adoption decisions--an urgent objective, given time lags and other considerations. Otherwise, it should probably emphasize functional contingencies among the three. This would yield information supporting both course improvement strategies and decisions as well as, possibly and more remotely, further curriculum design decisions. Indications of congruences between intentions (goals) and what occurs would seem to relate to decisions concerning curriculum design, and to dissemination and installation methods and practices.

It would appear that the payoffs for NSF of one or another form of absolute evaluation of MACOS hinge to an extent on how it views its role vis á vis the curriculum. If it wants to assure that the curriculum whose development it has sponsored receives maximum support and assurance of effective and continued use, then an absolute evaluation approach is probably the best vehicle. If it sees its role as spawning further curriculum projects in the social studies, then absolute evaluation may or may not be the most productive or appropriate for obtaining necessary information. It may well not bear on the decisions NSF wants to make. In either case, there will of necessity be long lead times. They can be reduced to some extent by interim reports. In the case of the first role, however, there must be special consideration of the feedback loops and feedback forms and formats between study and users.

Comparative Evaluations

"The conclusion seems obligatory that comparative evaluation, mediated or not, is the method of choice for evaluation problems."[15] We started this study with the presumption that comparative evaluation was the next logical and necessary step in evaluation of MACOS. Hanley's formative evaluations included comparison groups. Would not a useful further appraisal involve extension of the comparisons made in Hanley's studies?

The problem is, what are meaningful comparisons? The criterion of meaningfulness ought to be that the comparison leads to or is germane to decisions of general consequence. It is not clear in this instance what such decisions might be. To adopt MACOS instead of some other social studies curriculum? To launch another large-scale development project? To modify the MACOS curriculum? To use more multi-media curricula? To develop more inter-disciplinary curricula? To develop more task or goal oriented curricula? Perhaps the real issue is methodological. Can meaningful comparisons be made without actively manipulating some experimental variables? Scriven's engaging discussion of practical control group evaluation uses examples in which the evaluator has the power to manipulate some interesting independent variable. Hanley's comparisons were drawn from targets of opportunity, so to speak. Is it likely that the use of selected in vivo contrasts can ever be anything other than suggestive? This line of analysis quickly becomes an analog of the inquiry process students are being taught. The differences are the time scale,

---

[15] Scriven, Michael, op. cit. Scriven's position is, basically, that evaluation eventually is the judgment of worth or merit. This position is applicable to absolute evaluation, but it clearly gains in feasibility by comparative evaluation.

the presumed need to make practical decisions, and the more disparate relationships of those engaged in the process.

The basic issue of comparative evaluation in field settings is what really is being compared. Hanley, et. al., made comparisons of attitudes of children in MACOS and non-MACOS social studies courses, and of interview responses of MACOS and a limited number of non-MACOS teachers. Their other independent variable was type of school system. They did not attempt to compare social studies programs as such, nor did they attempt to measure cognitive or behavioral outcome variables directly. Their findings of different reactions of children within MACOS treatment groups suggest that there may well be pupil-program-pedagogy interactions of importance. But the restricted range of outcome variables explored across treatment groups makes it difficult to evaluate the significance of the findings within the MACOS groups.[16]

Comparative evaluations are notoriously treacherous, especially if they cannot employ true experimental control. Comparisons of intact groups observed in natural settings are subject to a variety of interpretational hazards when it comes to making attributive statements. Unfortunately, however, complete or even constrained or restricted randomization of assignments is unfeasible on a large scale at this point in time. (It might have been feasible during the development stage of MACOS.) We shall discuss explanatory experimental courses of action later, but in the present context we do not believe true experimental comparative evaluation is a feasible model. Nevertheless, MACOS has goals that are similar to the goals of other programs; or it seeks to stimulate skills, behavior, and attitudes, both in teachers and students, that may or could occur with other programs. Only comparative evaluation will help identify and clarify similarities and differences. If a comparative evaluation is to yield more information than is already available in Hanley's study and if it is not experimental, it should consider at

---

[16] These points are not at all intended to denigrate Hanley's extraordinarily thorough, informative, and creative formative studies.

least the following:

General Design Requirements

Most are similar to those for absolute evaluation.

1. Should include repeated observations of the same student, teacher, and school populations for at least three years. This is especially important with respect to MACOS teachers, who face the task of learning a new curriculum.

2. Should include a variety of prior, concurrent and subsequent organizations of instruction, educational orientations, and demographic settings, both for MACOS and non-MACOS groups. This obviously could easily result in a huge number of combinations of variables. It would probably be most appropriate to select clusters of schools and programs within geographically and demographically similar areas. This would permit a number of specific sub-comparisons, in addition to aggregate ones.

3. There should be a variety of specifically selected comparison social studies programs. That is, rather than MACOS versus non-MACOS, comparisons should be made, say, for fourth, fifth, or sixth grade students in program A, program B, program C, etc. Comparison groups should include the best or most innovative social studies programs available. The risks of this are obvious, and they serve to clarify the choices involved in comparative evaluations. One could test several curricula, including MACOS, under controlled and experimental conditions, using the best and most motivated teachers possible. This would lead to statements of high precision, but of uncertain application. Or one can do what is suggested here: seek performance measures for several contrasting curricula obtained under natural conditions. This will lead to statements of uncertain precision and greater applicability. For example, it should lead to statements of the following form:

> x percentage of MACOS children in the lower (middle, upper) third of the class in inner city fifth grades showed evidence in interviews of using a variety of

sources of information in gathering data pertaining to certain questions, compared with y percentage of similar children in non-MACOS program A (B,C, etc.).

Such statements will not explain anything; they will simply depict what one may expect to find on the average under various conditions of utilization. One reason for suggesting that observations be made in relation to the same schools and teachers to the extent possible over a period of several years is to obtain a measure of the stability of such statements over time. Statements of the sort suggested above can be amplified to include more conditions, or to refer to more specific sub-groups. They would also apply to follow-up observations with samples of children in subsequent years.

4. There should be a variety of outcome measures across programs. There are three basic problems here. One is that there should be observations or measures related to each of the seven MACOS goals. The second is that not all social studies programs necessarily share the same goals, or would accept the same evidence or criteria of goal behavior. The third is that content and concepts differ among programs. The measurement and instrumentation problem is central and probably poses the major dilemma and choice for continuing evaluation of MACOS, whether absolute or comparative.

There should be no serious problem in compiling or constructing generalized attitude scales and activities checklists for administration to students and teachers, regardless of program or curriculum. Similarly, there are scales that observers, including teachers, can use to rate the extent to which students demonstrate evidence of particular cognitive skills presumably related to inquiry processes, inductive thinking, and so on. There are classroom interaction checklists or scales that are independent of course content, including those used in Hanley's studies. Sets of non-content specific items could be used to obtain measures of performance

across programs. Some items or scales will vary in their application to different courses and classes. At least, the categories of behaviors or psychological functions presumably measured by different items may be more or less consistent with the categories of behaviors or functions developed by this or that course. But surely the reason for making comparisons is to obtain some indication of what MACOS children, for example, are likely to do in relation to what non-MACOS children do on the same measures.

Interviews could be developed to ascertain whether children who have had MACOS are more likely to give evidence of having applied what they learned, or maintained an interest germane to the course outside the classroom, than children in other social studies courses. There is, however, no convincing way that we can see of directly comparing concept or knowledge attainment, utilization, and retention across courses without showing equivalence of difficulty and/or functional equivalence of practice.

This argument does not apply to such MACOS goals, however operationalized, as encouraging "children to reflect upon their own experiences," or to "...legitimize the search."[17] Here it would seem to be precisely whether or not children in MACOS are likely to give more evidence of these (implied) behaviors than children in other social studies courses that is of interest. By the same token, if there are generalized skills, attitudes, and behaviors or processes that are stimulated and promoted in other social studies projects, it would be reasonable to ask to what extent they occur in MACOS children, intended or otherwise.

Choice of measures and instruments, be they paper and pencil tests,

---

17 Education Development Center. Man: A Course of Study Evaluation Strategies, op. cit., p. 12.

observer ratings, interviews with individuals or with groups, assignment of group projects, use of creative formats, transcripts of class discussions or whatever, is important on two counts. One is technical, the other cost. If there is the need for a variety of measures, and we believe there is, one is faced with the decision either to administer what could amount to substantial batteries to groups of children and/or classes, or to administer samples of items from different sets. The latter is the better choice, since with the possible exception of certain follow-up observations of children, we presume the comparative questions of interest apply to groups, not individuals. To our knowledge, however, there are not sufficient item pools available from which to draw for measures of applicable cognitive or process skills across several programs. It would be necessary to develop them, perhaps after the fashion of the National Assessment Program--a large and costly undertaking, or to patch together a battery of instruments for application at different points during the year and thereafter. The battery would probably have to be limited in scope to be at all attractive to participating schools, and to avoid extravagant administration and processing costs.

The choice, then, appears to be between 1) a large scale instrument development activity for measuring or observing development in children of cognitive (including imaginal), emotional and behavioral skills and orientations related to social studies, and 2) the use of a limited set of instruments of uncertain reliability and adequacy. The first would perhaps be of most value in the field of elementary social studies in the long run. It is difficult, however, to see how a large scale effort would be justified simply to permit, initially, comparative evaluation of MACOS. The latter is

feasible, although it will surely yield the more equivocal results.

As a final comment, it should be apparent that the preceding analysis applies primarily to the measurement of effects on children. Measures of outputs with respect to teachers, schools and communities seem less problematic, although here too theoretical and methodological problems can easily become severe and costly.

Specific Design Requirements

1. Selection of programs or curricula.
   Given the kinds of comparative statements we have suggested, it is desirable to draw samples of each comparison program from known populations; but there probably is not enough available information to do that. It would probably be necessary to sample among and within school districts on geographic and demographic bases. Even then, the task of engaging the cooperation of schools and teachers may be difficult, if not impossible in some cases. This approach does not preclude inclusion of cross-subject variations in settings; e.g., MACOS and some other curriculum in conjunction with and not in conjunction with other process-oriented curricula such as Science--A Process Approach. It would be more appropriate to study that sort of interaction separately, however, that is, in separate studies.

2. Size of the study
   This will probably turn out to depend on the size of the differences one would like to detect, and on the extent to which it is deemed desirable and feasible to make observations in classrooms or otherwise to collect data during a school year. It is our opinion that detailed observations of processes or transactions are neither feasible nor necessary for the form of comparative evaluation discussed here. We have assumed that the relationships of primary interest are between input and output. Sample sizes also must be planned with attrition in mind. Attrition here, as with absolute evalutaion, refers to pupils, teachers, and programs.

Implementation Mechanisms

There are no obvious differences in choices from those discussed earlier, with the following possible exceptions. If the study is to be very large (e.g., 1,000 teachers), there may well be some advantage in establishing a MACOS evaluation project agency or collaborating with other agencies. This would also be true if it were decided to pursue an extensive instrument development course.

Objectives and Payoffs

Results of comparative evaluations would seem to bear most directly on adoption and continuation decisions at the school district and, conceivably, state levels. They would also, logically, bear more directly on decisions concerning future curriculum design and development. Depending on the measures made and on the outcomes, it is possible that results would have a bearing on decisions about in-service training and support.

The payoffs for NSF would appear to be principally the fulfillment of a complete development cycle and the development of further scientific knowledge about process-oriented curricula in the social studies. Results of a comparative evaluation may be useful for future decisions about curriculum improvement projects. As noted earlier, however, lead times will be long, even with annual reports of results.

Should NSF elect to undertake major development of instrumentation for measurement in the social studies, the payoff could be an interesting indirect form of course improvement regardless of curriculum. It was our observation, admittedly very limited, that the measurement and evaluation of students' progress in MACOS, at least, was a source of major uncertainty or ambiguity to teachers, and to others, from one frame of reference. If instruments and guidelines were available, they could be of use. This is not

in the least to suggest that fine teachers are not fully capable of evaluating their work with their students. They are, however, dealing with functions which, despite the technology of behavioral objectives, are not well understood or explicated. They are dealing with enormously complex patterns of behavior. Possibly aids in recognizing, classifying, and evaluating them would be quite welcome.[18]

We have taken the position that a true experimental comparison on a large scale is not a feasible alternative. This in no way precludes the feasibility of one or more localized true experiments. We suggest, however, that while the internal validity of such experiments may be high, external validity is apt to be very uncertain unless the local experiments are repeated with a range of settings and populations. In that case, however, it is difficult to imagine being able to implement comparison curricula satisfactorily. Nevertheless, we do not wish to suggest that we consider true experimental comparative evaluations impossible. True experiments hold the greatest hope for yielding explanatory statements or hypotheses. We shall return to suggestions for experimental research and evaluation running parallel to a larger study following the next section (see Explanatory Studies, p. 70-71).

---

[18] The recently published booklet on evaluation techniques for MACOS (Man: A Course of Study Evaluation Strategies) should be of great help to many teachers in this respect.

Systemic Evaluations

We have grouped under the heading of systemic evaluations studies whose focus is one or more of the various models, channels, mechanisms, procedures, arrangements, or organizational consequences of the dissemination, adoption, and support of MACOS. There is no sharp boundary between systemic evaluation, thus conceived, and curriculum evaluations, absolute or comparative. Ultimate criterion measures of performance will, theoretically, be the same in all cases. Intermediate criteria of effectiveness and efficiency will differ, however. Consider the difference between the following questions:

How effective is the curriculum in "develop[ing] in youngsters a process of question-posing (the inquiry method)"? [19]

How effective is this or that training program in developing in teachers skill to promote the question-posing process in children?

In the first case an answer could be obtained with no reference to variations in training programs or models if care were taken to guard against or minimize sampling biases. In the second case, an answer could not be obtained without specifying measurements of training programs.

There are an infinite variety of evaluative questions that can be examined with reference to MACOS. Some or all in one way or another overlap with or impinge on areas of investigation already well established in different centers and universities. Our limited field observations left us with the conviction that three crucial functions in adoption and implementation of the course are selection, training, and quality control.

---

[19] Educational Development Center, Man: A Course of Study Evaluation Strategies, op. cit., p. 12.

There were wide variations in all functions, ranging from no formal processes to elaborate and continuous activities. In some cases in which there had been apparently very good training for very capable teachers, the teachers nevertheless indicated specific tasks for which they felt inadequately prepared. Examples of variations in the other two functions have been mentioned throughout this report. Our argument is that

1. the processes and arrangements by which the implementation of the MACOS curriculum comes about and is supported can be the subject of evaluation; and,
2. the use of MACOS in different settings can have systemic consequences that can be the subject of evaluation; and,
3. either or both subjects of evaluation may be as worthwhile in the long run as summative evaluation of the MACOS curriculum.

It is clear, as we have noted before, that systemic evaluations are not independent of curriculum evaluations. Indeed, a truly comprehensive curriculum evaluation would include all the components and dimensions of systemic evaluation. It would also probably be impossible to do in a form, time frame and budget that would provide widely useable results.

One of the basic problems in systemic evaluation, especially in appraising training and support models, is similar to the problem of evaluation of MACOS: what are appropriate criterion measures or dependent variables? what are the parameters of different training and support models? what are the interactions or differential sensitivities of methods and teachers? Methodologically, systemic evaluations will perforce differ depending on whether one focusses on input (dissemination, adoption, installation, training, support) or output (impacts or effects on teachers, schools, systems, communities, colleges, state agencies, etc), or on both.

It is worth noting that, in a way, input evaluation, in the above sense, is the task of identifying convergences. Its aim is to establish relationships that hold hope of maximizing successful use of the curriculum. Output systemic evaluation is the task of identifying divergences--the various consequences of adoption and utilization of the curriculum. Its aim is to establish possibilities, anticipated or planned, and unanticipated.

There is no question about the feasibility of evaluating the diffusion of MACOS, and its short term and long term systemic impacts. Periodic surveys would be one vehicle for accomplishing this. There are many questions, however, about the feasibility of evaluating the effectiveness of various training and support mechanisms and approaches. Such study faces the same problems of performance criteria as judging the effectiveness of the course with students. But it has the additional difficulties of measurement or observation of adults--teachers, teachers of teachers, support personnel, supervisors, etc. And it faces the further problem of diversity of training and support programs (potentially) unlike the curriculum which is a fixed and constant package by comparison. It must also take into account the resources and orientations of different schools, systems, and communities, as well as of the publisher, Curriculum Development Associates, unless National Science Foundation or some other agency continues to support training programs through the regional centers.

As with evaluation of the curriculum, it would be necessary to measure a variety of dependent variables, and a variety of methods of observation, including interviews, checklists, and classroom observation. Evaluation of MACOS training models makes most sense, of course, only if there continue to be training programs for the curriculum. If, for example, the regional centers continued to provide training institutes, an

evaluation of the different models and approaches employed by the centers could be made.[20]

The mechanisms or options for undertaking systemic evaluation are most likely those discussed in connection with absolute evaluation. Since cross-model study should employ the same measures, or samples of items from the same pool of items (as with curriculum evaluation), and since there are different theoretical orientations among the implementers of different training models, it would be appropriate to have the first phase of a systemic evaluation concerned with training start with a design conference. At such a conference, trainers as well as measurement and evaluation specialists, teachers, and supervisory personnel should have the task of identifying the variables, measures, and criteria that should be included or considered.

What is the payoff of systemic evaluation for NSF? As ever, it depends on the goals and roles that the Foundation sets for itself. If a goal is course improvement in the social studies through further curriculum development projects, it would seem that only output systemic evaluation would be germane (evaluation of impacts and effects of the curriculum on schools and, hopefully, teacher preparation institutions). If a goal is course improvement through system improvement, then a broader systemic evaluation is germane.

---

20 It is possible to set up systemic evaluation projects that include training evaluation. Thus, for several different training programs, one could set up several CMAS analysis projects in a number of school districts, coupled with follow-up empirical evaluation. See Merrisett, Irving, et.al., A Model for Analyzing Curriculum Materials and Classroom Transactions, ch. 8 in Dorothy McClure Fraser, ed., Social Studies Curriculum Development: Prospects and Problems. 39th Yearbook, NCSS, 1969.

Explanatory Studies: A Final Note

There are many explanatory studies that can and should be done with the MACOS Curriculum. These are the traditional province of experimental research. It is likely that such studies will appear increasingly, especially in the form of graduate dissertations. They will find their way uncertainly back into the public domain in annual reviews of research (e.g., in the National Council of Social Studies magazine Social Studies) and become at least incorporated into professional and technical writings. Since graduate dissertations are undoubtedly the most economic form of potentially good research, it is strongly suggested that National Science Foundation consider subsidizing some each year in specific evaluative support of curriculum projects such as MACOS, it if does not already do so.[21] Since a number of questions about MACOS and other curricula are answerable only on the basis of well-controlled research studies that need not be national in scope,[22] support for a number of pertinent research dissertations could be a programmatic alternative of considerable merit. The utility

---

21 Lest it be supposed that doctoral dissertations, for example, are unlikely to contribute to theory, knowledge, and practice, one should note the centrality of Griffin's graduate thesis in Metcalf's chapter, Research on Teaching the Social Studies (in Gage, N.L., ed., Handbook of Research on Teaching. Chicago: Rand McNally, 1963, ch. 17); or the number of dissertation references in the Review of Research in Social Studies, 1967; Vol.XXXII, NO. 6, Oct. 1968, by Cox, C. Benjamin, et al. Sixty-two of the eighty-three references were to the Dissertation Abstracts.

22 The adequacy of the experiment will always be contestable, especially if it is small and significant differences do not emerge. This can be addressed to some extent by asking the dissertation writer to note the size of the differences he could expect to detect at some level of probability, given the necessary parameters of his experiment. The issue with much individual research is not what magnitude of differences should be detected, but what magnitude can be detected with the resources available.

of this approach would be enhanced if NSF had or could assemble a pool of interesting hypotheses related to MACOS specifically, or to social studies curriculum improvement generally, to use as a guide in seeking and/or funding specific thesis projects.

There are obviously other vehicles for conducting research into MACOS, and we do not mean to exclude consideration of them. The essential point is that we have tried to make a distinction between further course evaluation and experimental research studies of many of the multitude of theoretical problems and issues related to the course and its use. The distinction, of course, blurs and disappears as evaluative effort becomes more concerned with explaining relationships, differences, associations, etc., as well as delineating them.

SUMMARY AND CONCLUDING COMMENTS

In the preceding sections we have tried to delineate issues and alternatives for further evaluation of MACOS. The purpose has been to suggest to the National Science Foundation some bases for choosing among alternatives. There are many choices, depending on one's concept of evaluation and on one's purposes in evaluating. We have analyzed alternatives from the point of view of steps the NSF might take. We do not suggest all possibilities have been exhausted, even from that perspective, although we think most major ones have been considered. If the problem were approached from the point of view of a local school, a school system, a state agency or an independent researcher, for example, other alternatives or rationales might be appropriate. We have not been able to provide a concise definition of trade-offs among alternatives, certainly not in a quantitative sense. We have, however, at least provided a start at depicting what trade-offs may be.

The main decision criteria we have suggested for choosing among alternative strategies have been: 1) the kinds of decisions to which different sorts of evaluation information seem to be related; and 2) the goals and roles NSF may have or want to have. It has been obvious that neither criterion uniquely differentiates alternatives. The first criterion especially does not differentiate alternatives, at least at the global and unspecified level at which we have considered it.

We have not considered costs as a criterion for choosing among alternative strategies. It is apparent that costs will depend mainly on the size and depth of study made within different alternatives (absolute, comparative and systemic evaluations), and on the mechanism chosen for

implementation. They will also depend on the extent to which development of new instrumentation is undertaken.

Our position has been that it will be possible to make meaningful and defensible evaluative statements using methods and instruments currently available or reasonably easily developed. This is not at all the same as saying there is no need for development of better instrumentation. There is especially need for better instrumentation with respect to measuring higher order cognitive structures and skills, such as hypothesis seeking and problem solving, or social learning, or affective orientations and skills. A recent restatement of the need was made by Coffman:

> "There is a mountain of unfinished business in this area [test developmnet] if we are to provide something more than a distorted view of a small aspect of the output as we do today through tests of cognitive skills and subject matter knowledge."[1]

Nonetheless, further evaluation of MACOS is possible without undertaking a major instrument development effort.

Our basic position is that there is no one measure of effectiveness of MACOS. Each of the course's goals can be given operational definitions; or criteria for classifying and recording the occurrence of goal related behavior can be developed. The goals may be the development of processes, as well as products. The goals may include the development or enhancement of dispositions, orientations, or attitudes. They may refer to children or to teachers. In any case one can establish criteria for identifying appropriate behavior, and conditions for observing it can be set up--interviews, classroom projects, classroom ratings, paper and pencil

---

[1] Coffman, William E. Concepts of Achievement and Proficiency. Proceedings of the 1969 Invitational Conference on Testing Problems. Educational Testing Service. November 1, 1969, 3-11.